

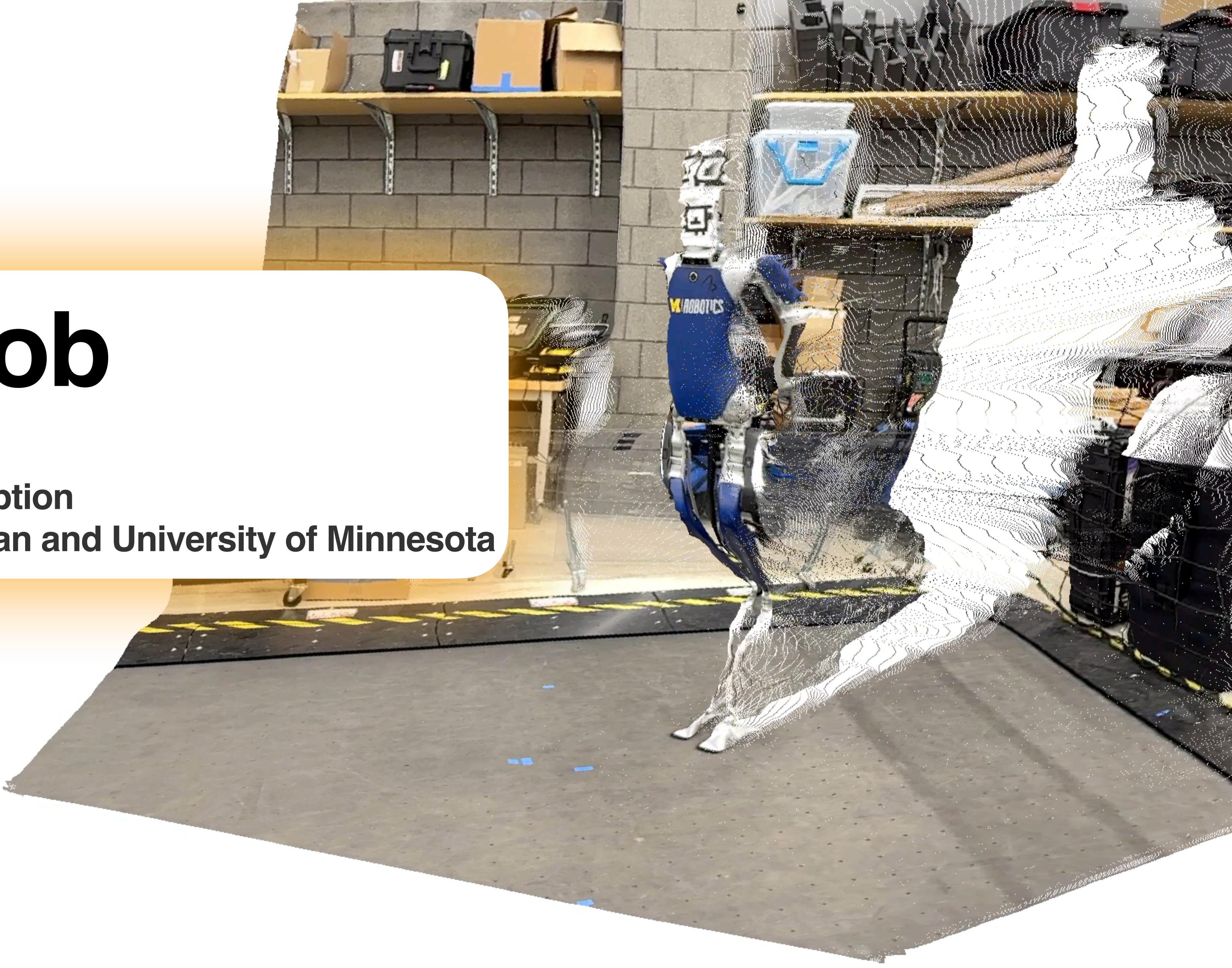
DR

DeepRob

Discussion 7

Prelude to 3D Perception

University of Michigan and University of Minnesota



Next Week: 3D Perception

- Seminar 1: RGB-D Architectures

1. [PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes](#), Xiang et al., 2018
2. [A Unified Framework for Multi-View Multi-Class Object Pose Estimation](#), Li et al., 2018
3. [PVN3D: A Deep Point-Wise 3D Keypoints Voting Network for 6DoF Pose Estimation](#), He et al., 2020
4. [Learning RGB-D Feature Embeddings for Unseen Object Instance Segmentation](#), Li et al., 2021

- Seminar 2: Point Cloud Processing

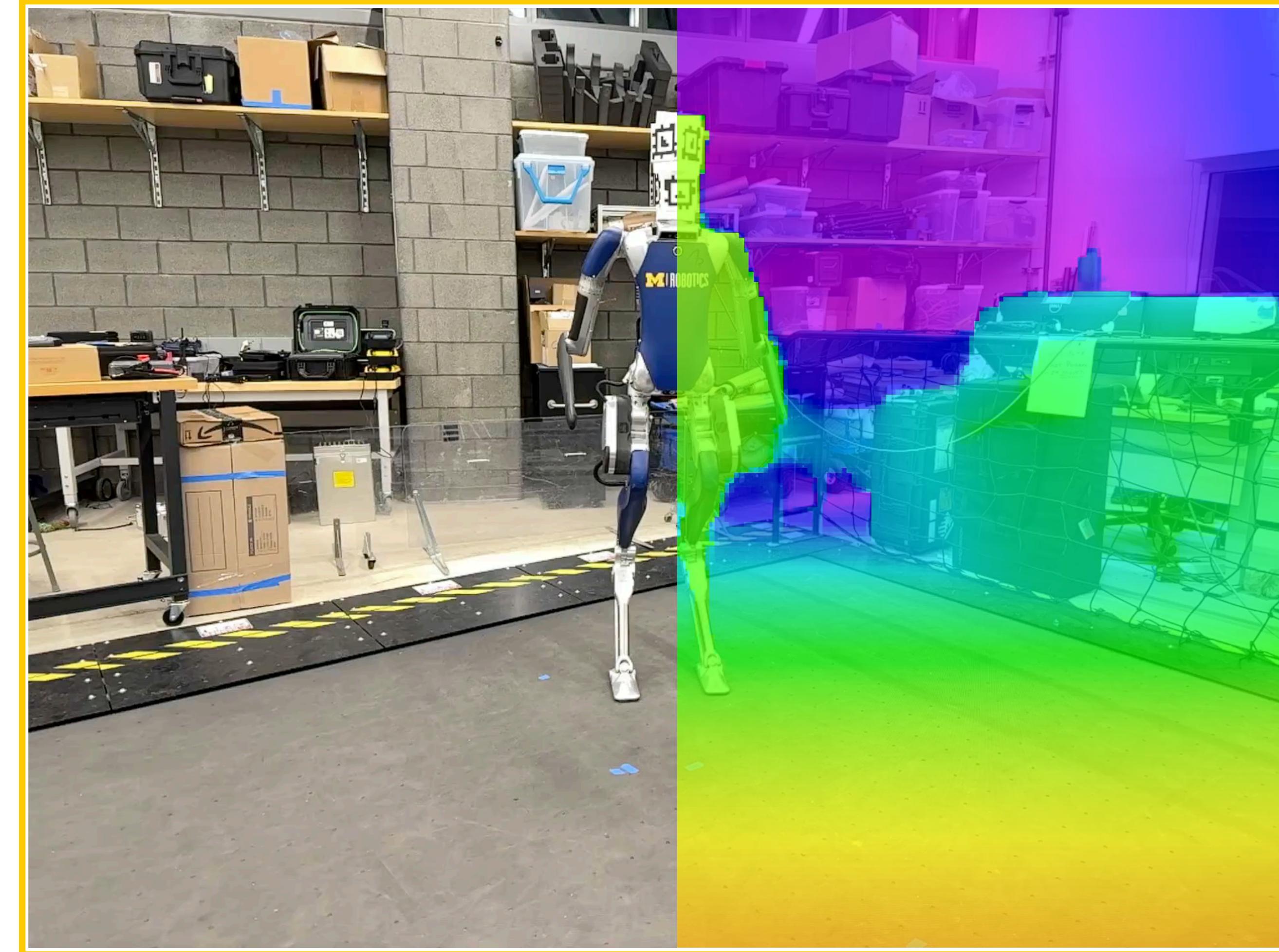
1. [PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation](#), Qi et al., 2017
2. [PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space](#), Qi et al., 2017
3. [PointFusion: Deep Sensor Fusion for 3D Bounding Box Estimation](#), Xu et al., 2018
4. [DenseFusion: 6D Object Pose Estimation by Iterative Dense Fusion](#), Wang et al., 2019

From 2D to 3D: Perception in a new dimension



“Video of WMT-TV switching from black and white to color broadcast”: [source](#) from CBS 2 Iowa

From 2D to 3D: Perception in a new dimension



Data courtesy of [Anthony Opiari](#), [Liz Olson](#), [Grant Gibson](#), and [Arden Knoll](#)

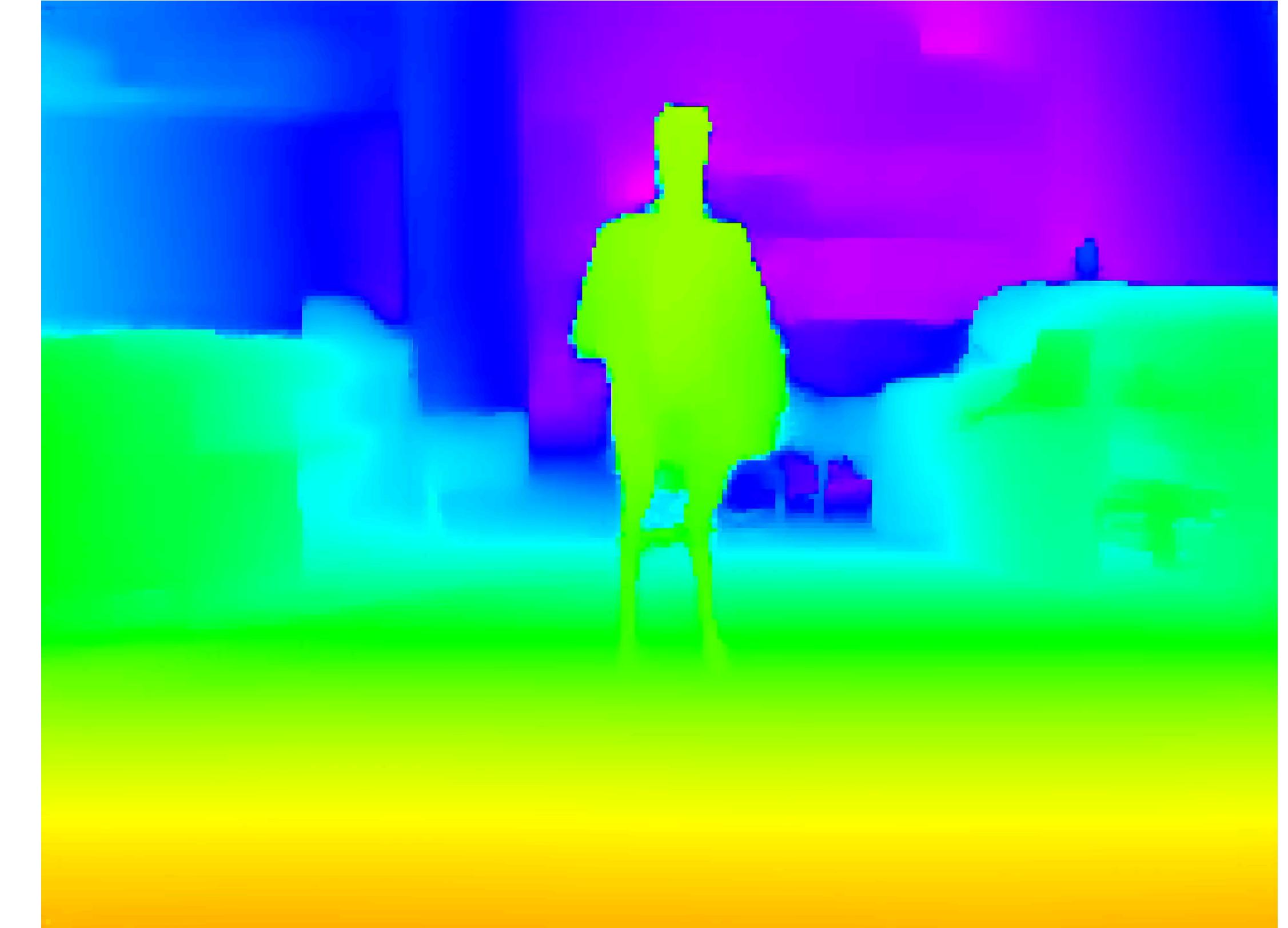
RGB-D Images

Color Image (RGB)



3 x 1920 x 1440

Depth Image (D)



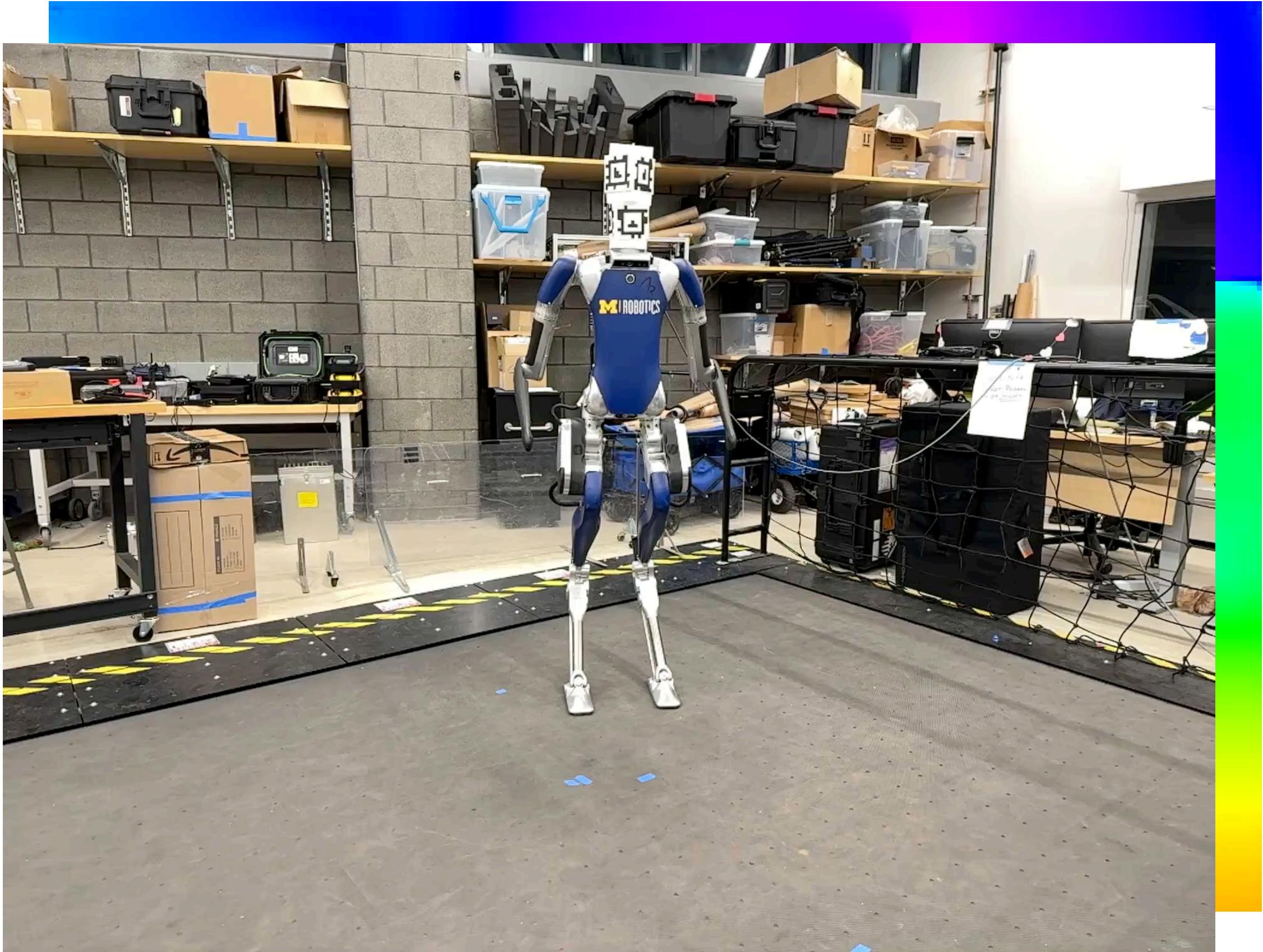
1 x 1920 x 1440



Data courtesy of [Anthony Oripipari](#), [Liz Olson](#), [Grant Gibson](#), and [Arden Knoll](#)

RGB-D Images

RGB-D Image



4 x 1920 x 1440

Deep Network

- 2D Object Detection
- 2D Semantic Segmentation
- 2D Instance Segmentation
- 3D Object Detection
- 3D Depth completion
- 3D Instance Segmentation
- 3D Reconstruction
- 3D Semantic Scene Completion
- 6DoF Pose estimation

RGB-D Sensors

Stereo cameras

1. Capture multiple images from displaced cameras
2. Find feature correspondences across images
3. Use known camera displacement to estimate depth image

Structured light sensors

1. Project known Infrared (IR) light pattern onto scene
2. Capture image in IR frequency
3. Use known pattern to compute depth image

Time of flight sensors

1. Produce infrared laser toward scene
2. Record time taken between transmission and response
3. Use known speed of light source to estimate depth image



Aside: Event Cameras

Consider time as the 3rd dimension

- Only record information as change occurs over time
- Better dynamic range than color cameras

CVPR 2021 Workshop on Event-based
Vision



Point Clouds

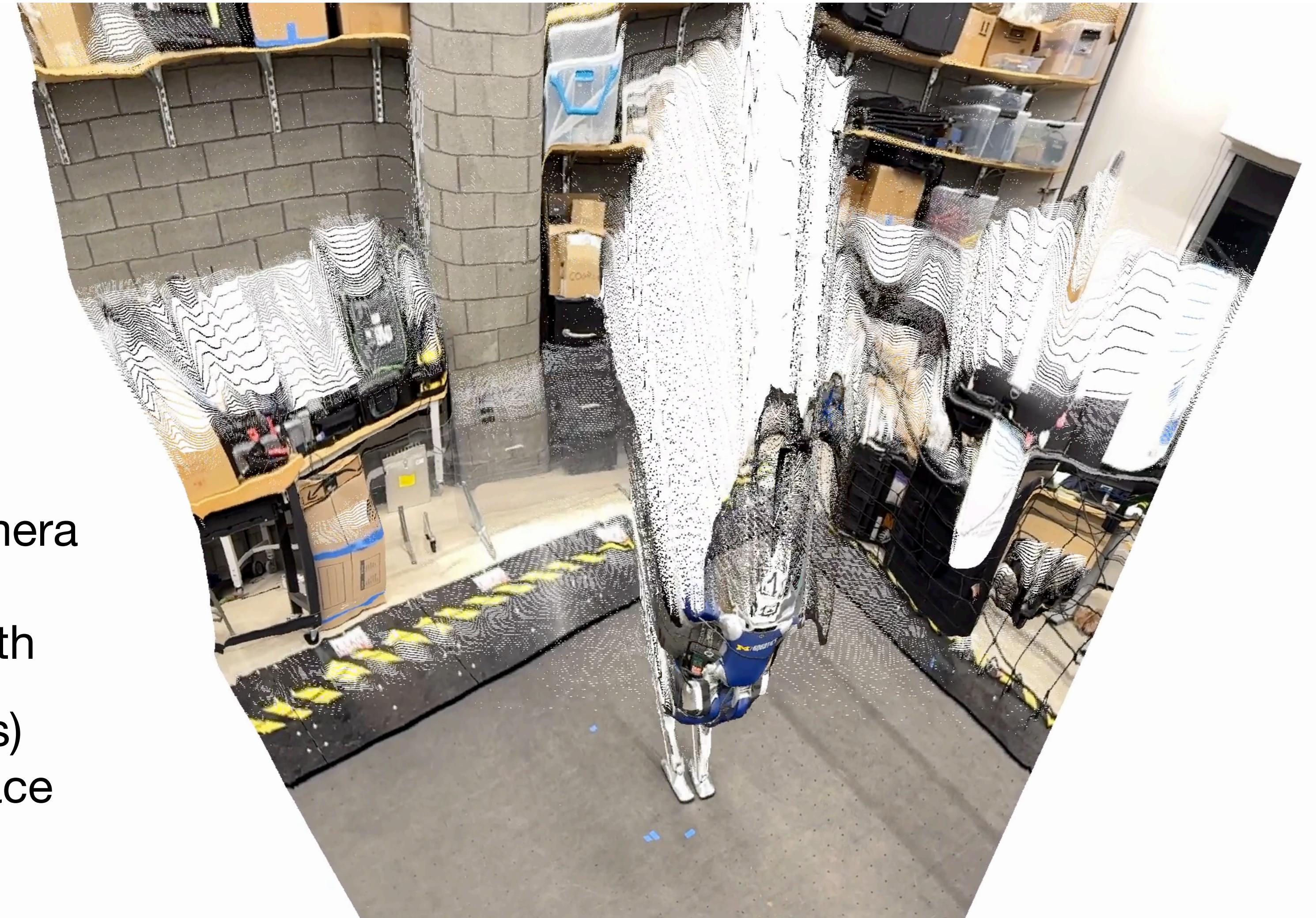
Point Cloud

A discrete collection of 3D points
Representation of 3D geometry
 $N \times 3$

Sources

Depth images with calibrated camera
transformed into point clouds by
projecting pixels along ray to depth

Object geometries (vertices, faces)
transformed by sampling on surface



Data courtesy of [Anthony Opiari](#), [Liz Olson](#), [Grant Gibson](#), and [Arden Knoll](#)

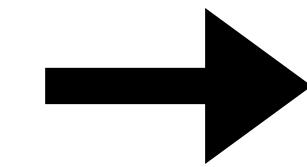
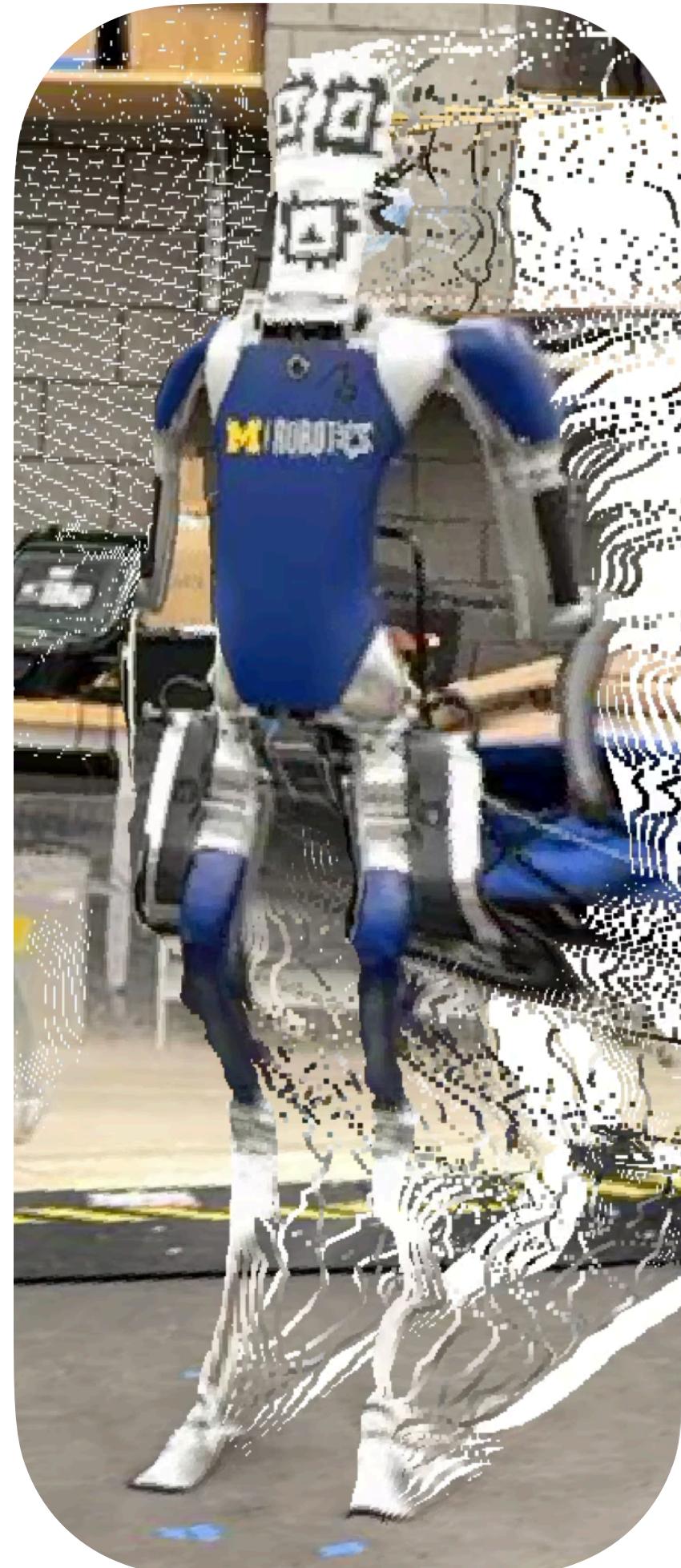
Recognition

Input

Point cloud: $N \times 3$

Output

Categorical label for object in point cloud



“Digit Robot”

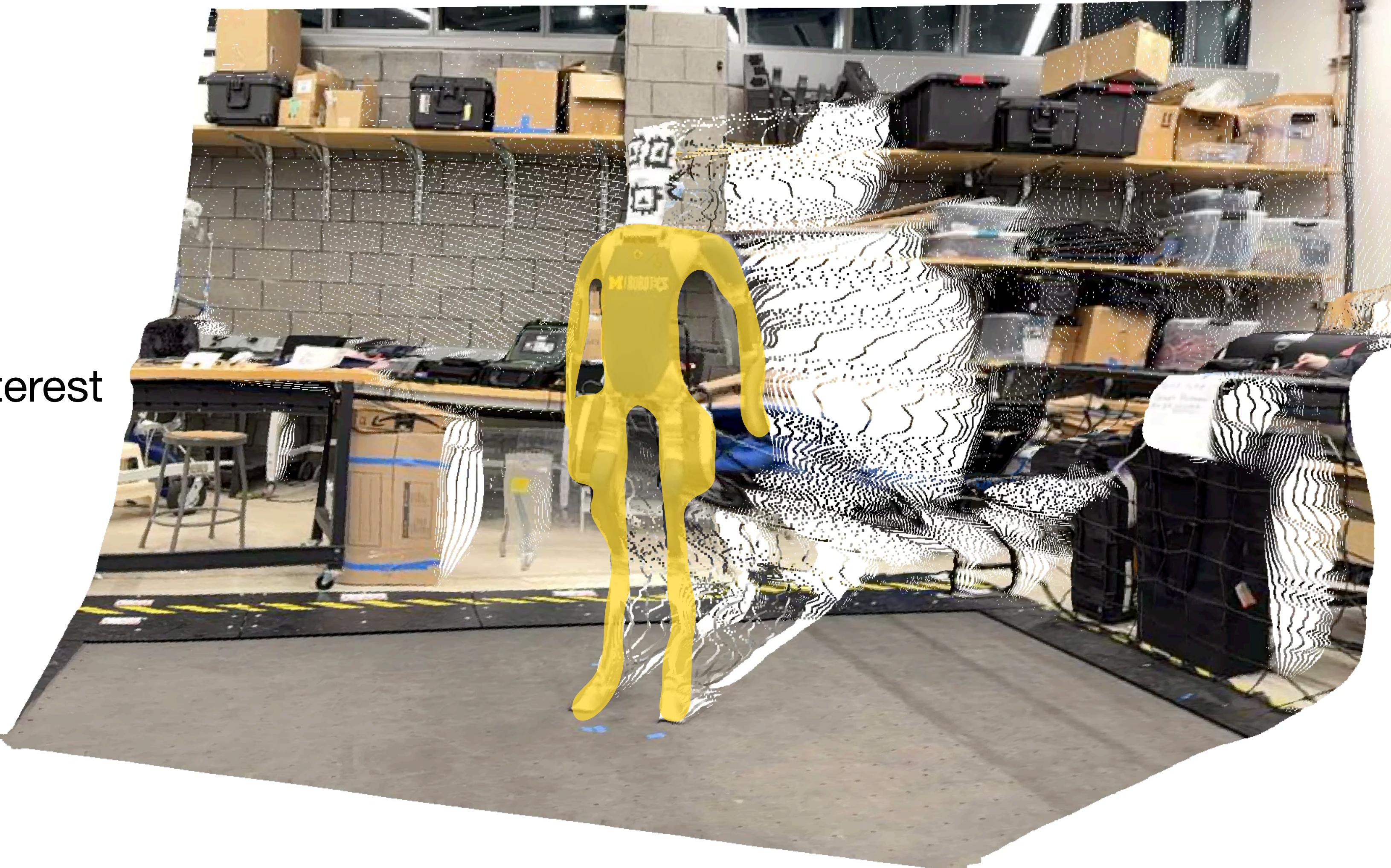
Segmentation

Input

Point cloud: $N \times 3$

Output

Categorical label for each point
corresponding to an object of interest
 $N \times (C+1)$



Registration

Input

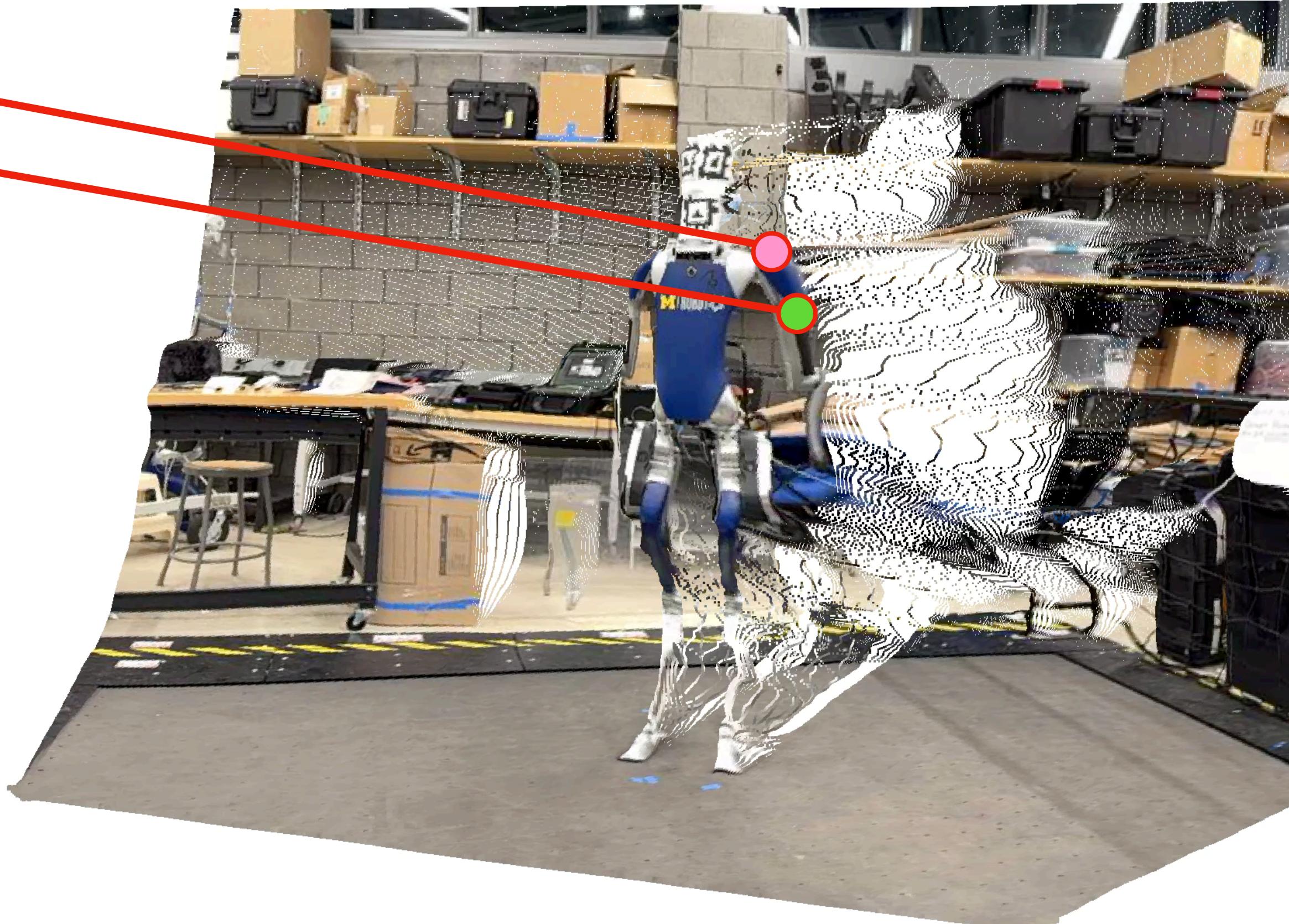
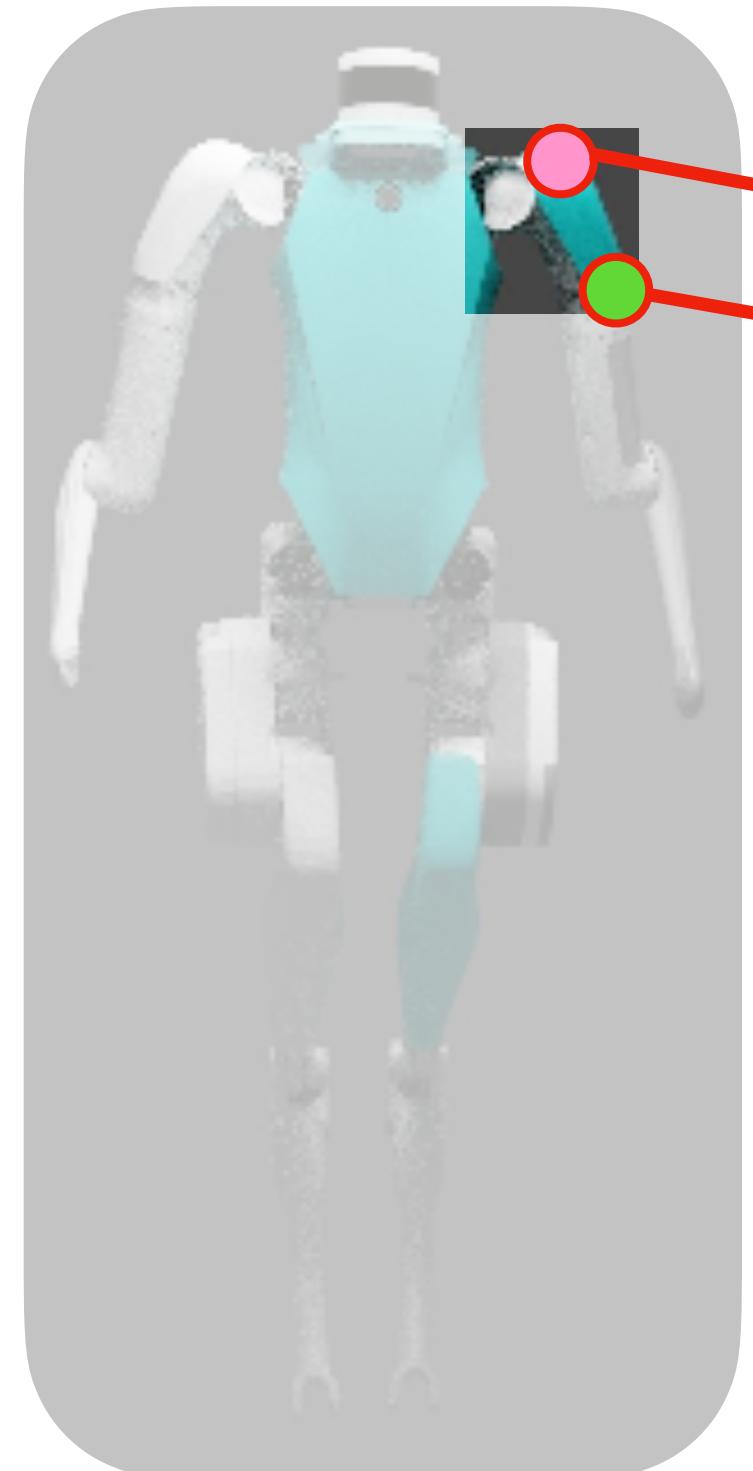
Object geometry (vertices, faces)

Point cloud: $N \times 3$

Output

Rigid body transformation for
geometry to point cloud frame

Pairwise point correspondences
between geometry and point cloud



Keypoint Detection

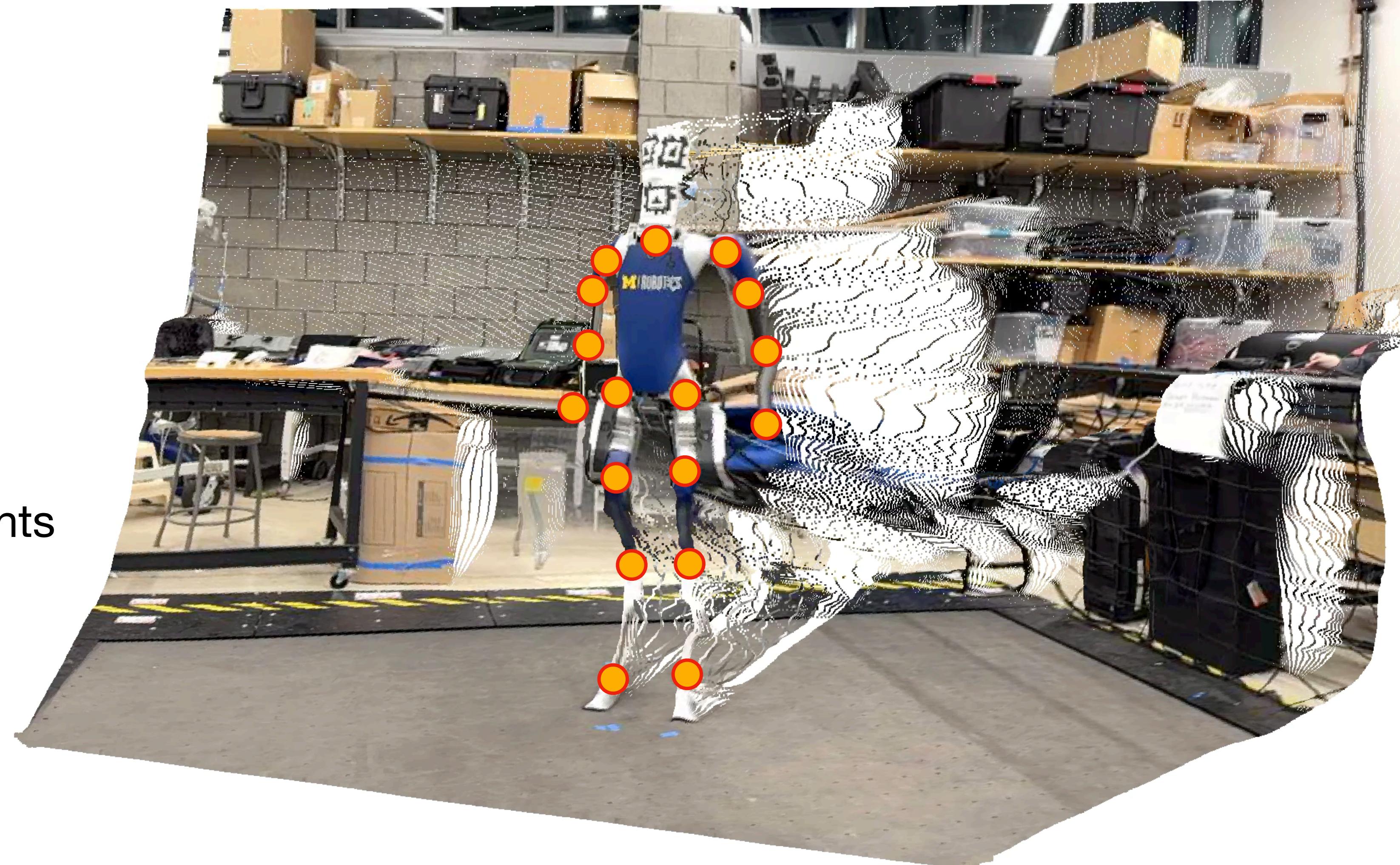
Input

Point Cloud: $N \times 3$

Output

3D Keypoint locations: 17×3

Corresponding to object link/joints



Data courtesy of [Anthony Opiari](#), [Liz Olson](#), [Grant Gibson](#), and [Arden Knoll](#)

Surface Reconstruction

Photogrammetry

Input Images



Data courtesy of [Anthony Opiari](#), Stanley Lewis

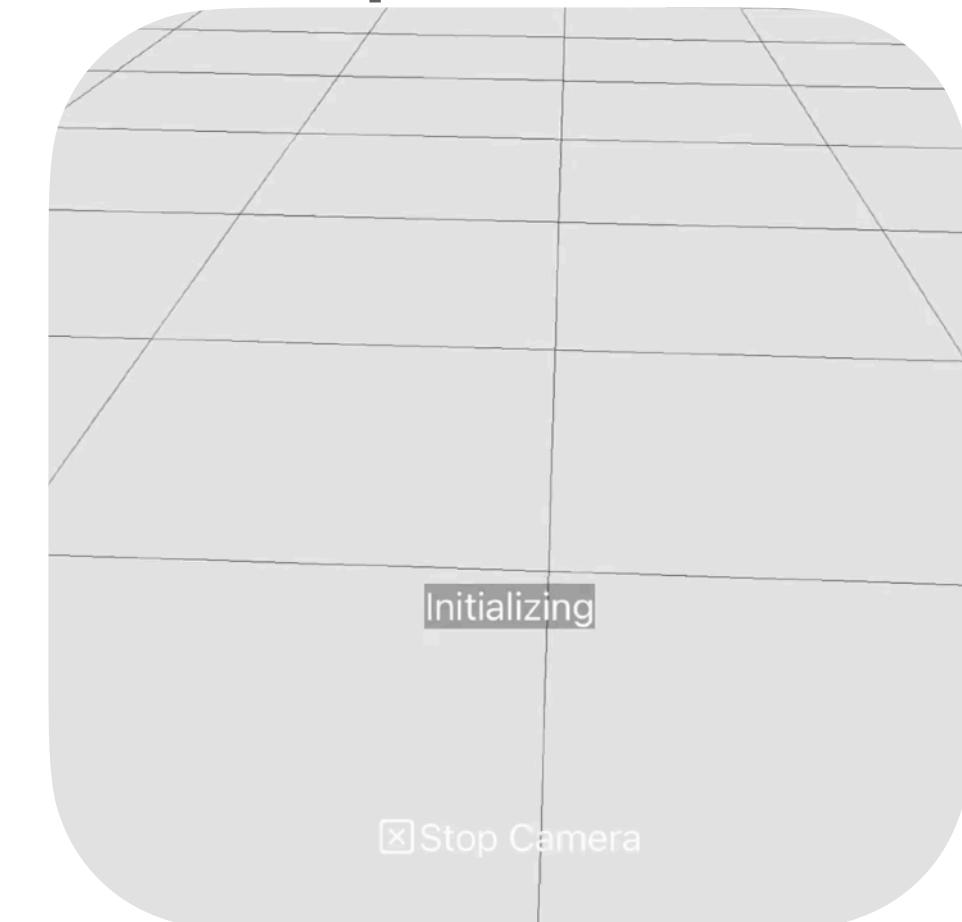
Output Geometry



Object model from [YCB Object and Model Set, Calli et al.](#)

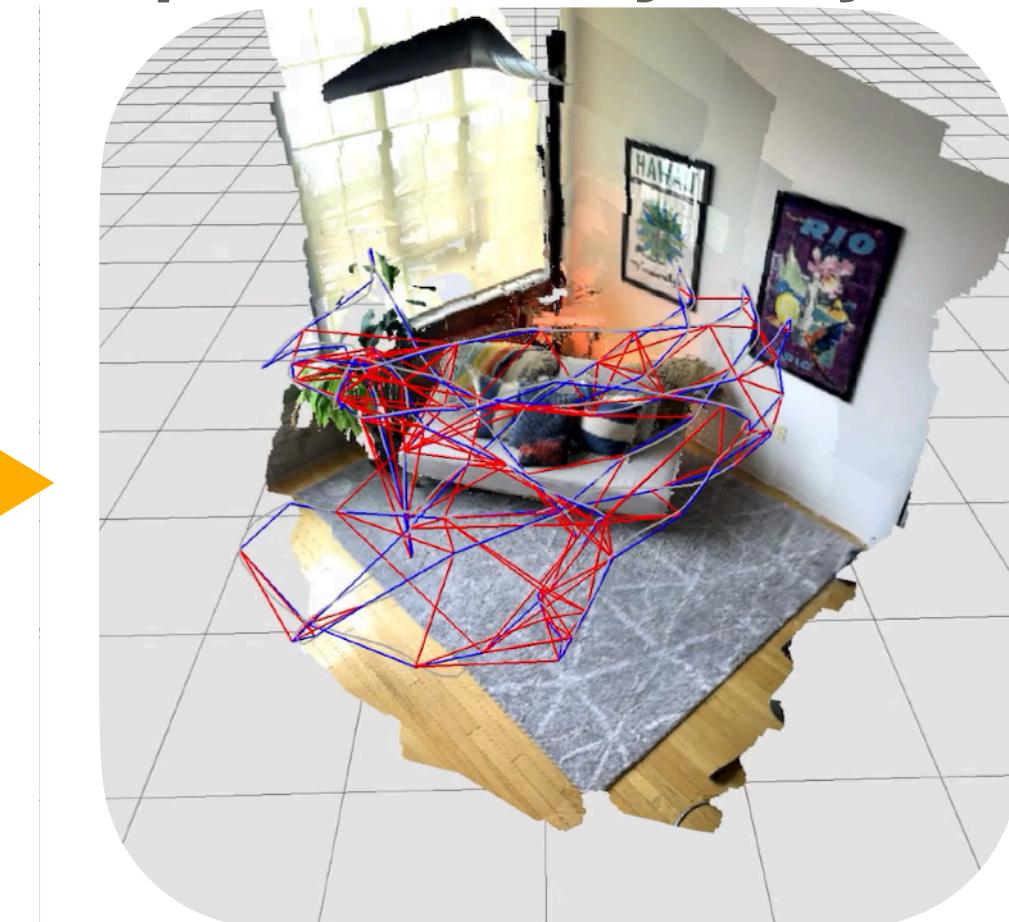
Mapping

Input Video



Data courtesy of [Anthony Opiari](#) using [RTAB-Map](#)

Output Geometry+Trajectory



Data courtesy of [Anthony Opiari](#) using [RTAB-Map](#)

3D Perception Resources

-  [Point Cloud Library \(pcl\)](#)
 - Large open source project for point cloud processing
 - Integrated with ROS
-  [PyTorch3D](#)
 - A Python library for deep learning with 3D data
 - Builds on top of PyTorch

DR

DeepRob

Discussion 7

Prelude to 3D Perception

University of Michigan and University of Minnesota

